TRANSMITTAL OF APPEAL BRIEF (Large Entity)

Docket No. 00240065aa

In Re Application Of: **L. Brocious, et al.**

Application No.	Filing Date	Examiner	Customer No.	Group Art Unit	Confirmation No.
09/670,646	9/27/00	M. Lerner	30743	2654	7162

Invention: **Explicitly Registering Markup Based on Verbal Commands and Exploiting Audio Context**

COMMISSIONER FOR PATENTS:

Transmitted herewith in triplicate is the Appeal Brief in this application, with respect to the Notice of Appeal filed on

The fee for filing this Appeal Brief is: **$500.00**

☐ A check in the amount of the fee is enclosed.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **09-0457**

☐ Payment by credit card. Form PTO-2038 is attached.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

Signature

Dated: **March 23, 2005**

Michael E. Whitham
Reg. No. 32,635

Whitham, Curtis & Christofferson, PC
11491 Sunset Hills Road - #340
Reston, VA 20190
703/787-9400
CUSTOMER NO. 30743

cc:

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to "Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450" [37 CFR 1.8(a)] on

_____.
(Date)

Signature of Person Mailing Correspondence

Typed or Printed Name of Person Mailing Correspondence

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re patent application of

Larry A. Brocious et al.

Serial No. 09/670,646

Filed September 27, 2000

For EXPLICITLY REGISTERING MARKUP BASED ON VERBAL COMMANDS AND EXPLOITING AUDIO CONTEXT

Group Art Unit 2654

Examiner Martin Lerner

Confirmation No. 7162

Commissioner for Patents
P..O. Box 1450
Alexandria, Virginia 22313-1450

APPELLANTS' BRIEF UNDER 37 C.F.R. § 1.192

This brief, which is filed herewith in triplicate, is in furtherance of the Notice of Appeal, filed February18, 2005.

This brief contains these items under the following headings and in the order set forth below, as required under 37 C.F.R. § 1.192(c):

I. REAL PARTY IN INTEREST

II. RELATED APPEALS AND INTERFERENCES

III. STATUS OF CLAIMS

IV. STATUS OF AMENDMENTS

V. SUMMARY OF INVENTION

VI. ISSUES

VII. GROUPING OF CLAIMS

VIII. ARGUMENTS

☐ ARGUMENT VIIIA. REJECTIONS UNDER 35 U.S.C. §112, FIRST PARAGRAPH

☐ ARGUMENT VIIIB. REJECTIONS UNDER 35 U.S.C. §112, SECOND PARAGRAPH

☑ ARGUMENT VIIIC. REJECTIONS UNDER 35 U.S.C. §102

☑ ARGUMENT VIIID. REJECTIONS UNDER 35 U.S.C. §103

☑ ARGUMENT VIIIE. REJECTION OTHER THAN 35 U.S.C. §§102, 103 AND 112

IX. APPENDIX OF CLAIMS INVOLVED IN THE APPEAL

X. OTHER MATERIALS THAT APPELLANT CONSIDERS NECESSARY OR DESIRABLE

I. REAL PARTY IN INTEREST

The real party in interest in the appeal is:

☐ the party named in the caption of this brief.

☑ the following party:

International Business Machines Corporation of Armonk, New York.

II. RELATED APPEALS AND INTERFERENCES

With respect to other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal:

☑ there are no such appeals or interferences.

☐ these are as follows:

III. STATUS OF CLAIMS

The status of the claims in this application is as follows:

A. Total number of claims in Application

The claims in the application are: Claims 1-21, totaling 21 claims

B. Status of all the claims:

1. Claims cancelled: None
2. Claims withdrawn from consideration but not cancelled: None
3. Claims pending: Claims 1-21
4. Claims allowed: None
5. Claims rejected: Claims 1-2, 11-13
6. Claims objected to: 3-10, 14-21

C. Claims on Appeal.

The claims on appeal are: Claims 1-2, 11-13

IV. STATUS OF AMENDMENTS

The status of amendments filed subsequent to the final rejection is as follows: There are no after-final amendments.

V. SUMMARY OF INVENTION

The claimed invention provides an extension to the HyperText Markup Language (HTML) allowing a user to employ context-sensitive audio commands to tell a browser what to present and what options are available for interaction with an application for which audio commands have been enabled. The claimed invention enables voice commands needed by an application, registers such commands with a speech engine, and provides an audio context for page-scope commands by adding a context option to make the page more flexible and usable. The invention thus enables a browser to respond to visual or verbal commands, or a combination thereof, by identifying what action will be taken based on the commands.

According to the prior art, applications, browsers, and speech engines are tightly linked together in a manner that prevents one application from working with multiple browsers or speech engines. As a result, current implementations have devices that will read aloud the words on a page but which require input to be entered either by keyboard or by an elaborate method such as where a user must proceed letter-by-letter using code words for letters of the alphabet, like "Alpha" for "A."

It is an object of the claimed invention to allow applications to register specific commands that will cause a browser to take an action based on the current audio context of the browser. It is a further object of the claimed invention to have a browser take an action based on current audio context and a word or words currently being spoken by a user. It is yet another object of the claimed invention to allow one application to work with multiple browsers and speech engines.

The claimed invention provides a generic way of encoding information needed by an application to register voice commands and enable the speech engine. This is done by introducing new HTML statements with the keyword META_VERBALCMD, which list the recognized/registered speech commands and what each one will do. This applies to commands that affect a whole PAGE in scope, like the "help" or "refresh" command. No matter where a user is on the page or what the user is doing, these commands work the

same and issue the same URL command to the user just as if the user had physically clicked on the HELP or REFRESH buttons on the screen.

The claimed invention further provides a sense of audio context. The context of a page changes as the audio presentation of the page progresses. The claimed invention adds the ability to alter the action based on the current audio context by adding the CONTEXT option to the META_VERBALCMD statements.

To take one possible example *inter alia*, the application may be a trip planner installed in an automobile and may be enabled to speak directions while displaying a map. A spoken command such as "repeat" may be employed to cause the application to speak the whole page of directions from the beginning. According to the claimed invention, however, it is possible to specify CONTEXT= "OPTIONAL" so that the browser may provide the application with a context to enable the application to tailor its response to the spoken command "repeat." Thus, if the user is listening to a direction at the time he or she speaks the command "repeat," the application would apply the command to the context and repeat the particular direction. If, however, the user is not listening to data from the application at the time she or she speaks the command "repeat" (*i.e.*, there is no current CONTEXT), the application would apply the command in the absence of context and speak the whole page of directions from the beginning.

Some spoken commands may be specified as CONTEXT="REQUIRED" instead of CONTEXT= "OPTIONAL". To take one example *inter alia*, a person may be reviewing email in an audio mode while driving. While an email application is reading aloud the topic of an email message or the name of the sender, a command such as "open" spoken by the user may cause the application to open and read aloud the contents of the message. According to the claimed invention, the performance of such an application could be improved by specifying CONTEXT="REQUIRED" to instruct the browser to recognize the spoken word "open" as a command only when there is an appropriate context recognized by the application at the time the word is spoken. If no such context is present when the word "open" is spoken, the word will not be recognized as a command. Thus, by way of example and not limitation, a user arriving at a rest stop

may speak the command "stop reading" to stop reviewing email. Such user may then tell passengers, "You can open the door now and get out," without causing the email application to interpret the word "open" as a command to open an email message. This would occur because of the absence of an appropriate CONTEXT under circumstances in which CONTEXT="REQUIRED" has been specified.

VI. ISSUES

The sole issue presented in this Appeal is whether Claims 1-2 and 11-13 are anticipated by U.S. Patent No. 5,732,216 to Logan et al.

VII. GROUPING OF CLAIMS

As noted above, claims disclosing a system and method for providing context based verbal commands to a multi-modal browser in which an audio context must be established to enable voice commands associated with the audio context to be recognized have been identified as being allowable (see claims 3-10 and 14-21), and therefore action with respect to these claims is not on appeal. Of the remaining claims, the claims are grouped as follows:

Claim Group 1. Claims 1, 11 and 12

Claim Group 2. Claims 2 and 13

Reasons as to why the grouped claims are separately patentable are included in the arguments.

ARGUMENT VIIIA. REJECTIONS UNDER 35 U.S.C. §112, FIRST PARAGRAPH

There are no rejections under 35 U.S.C. §112, first paragraph.

ARGUMENT VIIIB. REJECTIONS UNDER 35 U.S.C. §112, SECOND PARAGRAPH

There are no rejections under 35 U.S.C. §112, second paragraph.

ARGUMENT VIIIC. REJECTIONS UNDER 35 U.S.C. §102

Pursuant to an office action dated November 16, 2004 (the "Final Rejection"), Claims 1-2 and 11-13 were erroneously rejected under 35 U.S.C. § 102(b) as anticipated by U.S. Patent No. 5, 732,216 to Logan et al. Applicants respectfully submit that Claims 1-2 and 11-13 are not anticipated by Logan et al. because, among other considerations, the "context based" features of the claimed invention enable commands to be registered at different times during a documents' audible presentation, and to permit commands to have different meanings at different times depending on the context. The disclosure of Logan et al. does not address context based commands in any way.

The Examiner has argued that Applicants' position as to such "context based" features reads the specification of the claimed invention into the claims. (Final Rejection at 6) It is evident by the repeated use of the words "context based" in the claims, however, that the claims expressly disclose context-based features. Because Logan et al. does not teach or disclose such context-based features, the rejected are not anticipated by the reference. In making this argument, the Examiner applied a dictionary definition of "register" (Final Rejection at 8) in order to avoid the application of the term as it is used in the specification. (Specification, page 8, line 7; page 13, line 17)

The conclusion that Logan et al. does not anticipate the claimed invention is not surprising or extraordinary in any way, since the invention of Logan et al. concerns an audio messaging system, while the claimed invention concerns a system and method for providing context based verbal commands to a multi-modal browser.

CLAIM GROUP 1

Claim Group 1 (Claims 1, 11 and 12) is drawn to a system and method for providing context based verbal commands to a multi-modal browser. These claims stand rejected under 35 U.S.C. § 102(b) as anticipated by Logan et al. The claims of Claim Group 1 are distinct, and separately patentable, from the claims of Claim Group 2. Notably, for example, Claim Group 2 requires accessing different URLs, while Claim

Group 1 does not.

Claim 1, which may be taken as exemplary of Claim Group 1, is drawn to a system for providing context based verbal commands to a multi-modal browser, comprising:

> a context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user;
>
> a store for storing a current context of the audio queue; and
>
> a speech recognition engine for recognizing and registering voice commands, wherein said speech recognition engine compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison.

Claims 1 and 12

With regard to Claim 1, the Examiner erroneously found that "*Logan et al.* discloses a system for controlling an audio controller" (Final Rejection at 2) and that the invention disclosed by Logan et al. is equivalent to Claim 1's "system for providing context based verbal commands to a multi-modal browser." The method of Claim 12 has been rejected on substantially the same basis. (*See* Final Rejection at 4) Applicants respectfully submit that this is in error.

Claim 1 and Claim 12 enable audio commands to be obtained from an input markup and allow users to speak such commands to bring about an action. The commands thus registered are dynamic in nature and need not be the same for every page, a feature not disclosed or taught by Logan et al. Even though the specification of Logan et al. mentions the use for audio commands for navigating a system, there does not appear to be anything to indicate that Logan et al. ever recognized problems relating to how to get a browser to take an action based on the current audio context of the browser. By contrast, Claims 1 and 12 are directed to a system and a method "for providing context based verbal commands to a multi-modal browser," which is not accomplished or discussed by Logan et al. Nor does there appear to be anything in the disclosure of

Logan et al. to anticipate "registering voice commands, wherein said speech recognition engine compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison," as in Claims 1 and 12. Not only does Logan et al. not appear to contemplate registration of commands, such commands appear to be of a fixed nature in Logan et al., supporting only a standard set of navigation keywords designed to supplement conventional automobile radio, tape of CD controls:

> The ability to navigate the program using only audio prompts and/or small number of buttons for a user interface make the playback system which utilizes these features of the invention particularly attractive for use by automobile drivers, who can select their program content much more effectively and with less drive distraction than currently possible with a conventional automobile radio, tape or CD player.

(Logan et al., column 35, lines 48-55)

The invention of Logan et al. is, therefore, not context-sensitive as in Claims 1 and 12. Applicants respectfully submit that the Examiner's finding that Claims 1 and 12 are anticipated by Logan et al. is based on a misapprehension of the reference, the claimed invention, or both.

In finding Claim 1 to be anticipated by Logan et al., the Examiner has relied extensively on Figure 5 from the disclosure of Logan et al. However, nothing in Figure 5 of Logan et al. refers to a "multi-modal browser," and, because Figure 5 makes no provision for context sensitivity, there is nothing to anticipate a "context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user," "a store for storing a current context of the audio queue," "a speech recognition engine [which] compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison," or the equivalent of any of those features.

Similarly, in finding Claim 12 to be anticipated by Logan et al., the Examiner has relied on Figure 5, discussed above, and also on Figure 1 from the disclosure of Logan et

al. However, nothing in Figures 1 and 5 of Logan et al. refers to a "computer implemented method for providing context based verbal commands," "building a context based audio queue based on the contents of markup language page being audibly read by the multi-modal browser," "storing a current context of the audio queue," "recognizing and registering voice commands, wherein the current audio context is compared with a voice command," "causing the multi-modal browser to perform an action based on the comparison," or the equivalent of any of those features.

Just as Figures 1 and 5 of Logan et al. do not anticipate Claim 12, the various portions of the specification of Logan et al. cited by the Examiner do not anticipate Claim 12, either. For example, the Examiner has relied on the same passages to show that Logan et al. discloses both "a context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user" (Final Rejection at 2), as in Claim 1, and "building a context-based audio queue based on the contents of markup language page being audibly read by the multi-modal browser to a user" (Final Rejection at 4), as in Claim 12:

> As contemplated by the invention, information which is available in text form from news sources, libraries, etc. may be converted to compressed audio form either by human readers or by conventional speech synthesis. If speech synthesis is used, the conversion of text to speech is preferably performed at the client station 103 by the player. In this way, text information alone may be rapidly downloaded from the server 101 since it requires much less data than equivalent compressed audio files, and the downloaded text further provides the user with ready access to a transcript of voice presentations. In other cases, where it is important to capture the quality and authenticity of the original analog speech signals, a text transcript file which collaterally accompanies a compressed voice audio file may be stored in the database 133 from which a transcript may be made available to the user upon request.

(Logan et al., column 5, lines 16-45); as well as

As hereinafter described in connection with FIG. 5, each voice or text program segment preferably includes a sequencing file which contains the identification of highlighted passages and hypertext anchors within the program content. This sequencing file may further contain references to image files and the start and ending offset locations in the audio presentation when each image display should begin and end. In this way, the image presentation may be synchronized with the audio programming to provide coherent multimedia programming.

(Logan et al., column 5, lines 6-15); and

In addition, the structured program files may advantageously contain, where appropriate, "hyperlink" passages, which may take the form of announced cross references to other materials, or sentences or phrases which describe related information contained elsewhere in the download compilation but which do not follow immediately in the sequence. In order to alert the listener to the fact that a sentence or passage is a hyperlink to other information which is out of the normal playback sequence, an audible cue may advantageously proceed, accompany, or immediately follow the passage in the normal playback which identifies the character of the hyperlinked material. Using the terminology typically employed to described hypertext, the normal programming sequence includes "anchor" passages which are identified by an audible cue signal of some type and are further associated with a reference to hyperlinked material to which the playback may jump upon the listener's request. Hyperlinked material, like all other programming, is advantageously preceded with a topic description and, if the hyperlinked material is a narrative, it should begin with a summary paragraph, followed by increasing detail.

A hyperlink may be directed to a program segment which is not present in the current selections list. In that case, the Link variable

contains a negative number to distinguish it from references to a particular Selection_Record, and is interpreted as the negative of a ProgramID number. If the referenced ProgramID is available in the player's mass storage system, it may be fetched an played and, upon its conclusion, an automatic return is made to the original sequence. If the referenced ProgramID does not refer to a locally stored record, the listener is informed that it is currently unavailable, but will be included in the next download for the next session.

In addition to having means for accepting a user command to execute a jump to the hypertext material, the player also advantageously includes a mechanism (special key or voice command response) which, when activated, causes a "return" to be made to the playing sequence at the point of the original anchor from which the hyperlink was performed. In this way, a listener may listen to as much or as little of the linked information as desired, retaining the ability to return to the original. Just as computer subroutines may be nested by saving the return addresses of a calling instruction in a stack mechanism, a hyperlink may be executed from within a hyperlinked narrative, and so on, with the listener retaining the ability to execute a like (Logan et al., column 30, lines 20-66)

The portions of the disclosure of Logan et al. cited by the Examiner do not refer to a context-based audio queue, especially given the fact that Logan et al. does not address matters involving context such as are addressed by the claimed invention. Nor is there anything in the cited portions of Logan et al. which anticipates the use of context sensitivity, either in connection with a multi-modal browser or otherwise.

Similarly, the Examiner has relied on the following passages to show that Logan et al. discloses "a speech recognition engine for recognizing and registering voice commands, wherein said speech recognition means compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison." (Final Rejection at 3):

The player 103 further includes a sound card 110 which receives audio input from a microphone input device 111 for accepting voice dictation and commands from a user and which delivers audio output to a speaker 113 in order to supply audio information to the user.

(Logan et al., column 3, lines 32-37); as well as

User Playback Controls

The player mechanism seen at 103 includes both a keyboard and a microphone for accepting keyed or voice commands respectively which control the playback mechanism. As indicated at 261, the receipt of a command, which may interrupt the playback of the current selection, and the character of the command is evaluated at 262 to select one of six different types of functions.

(Logan et al., column 12, lines 50-58); and

Whenever the user issues a "Go" command (seen at 265 in FIG. 3), the player will execute a hyperlink jump to the location indicated by the last "L" record in the selection file. When the jump is made, the location in the "L" record is inserted into the CurrentPlay register 353 after the previous contents of the CurrentPlay register are saved in (pushed into) a zero-based stack 390 at the stack cell location specified by the contents of a StackPtr register 392, which is then incremented. Whenever the listener issues a "Return" command, the previously pushed selection file record location is popped from the stack 390 and returned to the CurrentPlay register 353, and the StackPtr register 392 is decremented. A "Return" command issued when StackPtr=zero (indicating an empty stack) produces no effect.

(Logan et al., column 35, lines 1-15).

While the cited portions of the disclosure of Logan et al. contemplate the use of speech recognition as a general matter, there is nothing to anticipate the possibility of context-sensitive uses of speech recognition, which is characteristic of Claim 1.

Applicants respectfully submit that the disclosure of Logan et al. does not

anticipate Claims 1 or 12 of the claimed invention.

Claim 11

With regard to Claim 11, which depends from Claim 1, the Examiner found that "*Logan et al.* discloses the host server stores web page data 141 by means of an HTNL interface . . . HTML web server 129 presents HTML program selection forms . . . narrative text is presented in the interactive, multimedia format expressed in the first instance using essentially conventional hypertext markup language." (Final Rejection at 5) Applicants respectfully submit that the Examiner erred. the rejection of Claim 11.

In finding Claim 11 to be anticipated by Logan et al., the Examiner has relied on Figure 1, discussed above, and Figure 7 from the disclosure of Logan et al. Nothing in Figures 1 and 7 of Logan et al. discloses the substance Claim 1, including context based features, while adding "wherein the page being audibly read is a markup language page." (Claim 11) Just as Figures 1 and 7 of Logan et al. do not anticipate Claim 11, the various portions of the specification of Logan et al. cited by the Examiner also do not anticipate Claim 11. The Examiner has relied on the following passages to show that Logan et al. discloses "the host server stores web page data 141 by means of an HTML interface." (Final Rejection at 5):

> The host server 101 further stores web page data 141 which is made available to the player 103 by means of the HTML interface 128. The host server 101 additionally stores and maintains a user data and usage log database indicated

(Logan et al., column 5, lines 32-35) The cited passage does not anticipate Claim 11 because it does not teach the substance Claim 1, discussed above, while adding the limitation "wherein the page being audibly read is a markup language page."

In addition, the Examiner has relied on the following portion of the disclosure of Logan et al. to show that "HTML web server 129 presents HTML program selection forms." (Final Rejection at 5):

> In addition to the downloaded catalog of available items which may be viewed by the subscriber from the available downloaded information, the

user may re-establish an Internet connection to the HTML web server 129 which presents HTML program selection and search request forms, enabling the subscriber to locate remotely stored programming which may be of particular interest to the subscriber. When such programs are selected in the HTML session, the user's additional preferences and selections may be posted into the user data file 143 and the identification of the needed files may be passed to the client/player 103 for inclusion in the next download request.

(Logan et al., column 8, lines 48-60) Again, the cited passage does not anticipate Claim 11 because it does not teach the substance Claim 1, discussed above, while adding the limitation "wherein the page being audibly read is a markup language page."

The Examiner also relied on the following portion of the disclosure of Logan et al. to show that "narrative text is presented in the interactive, multimedia format expressed in the first instance using essentially conventional hypertext markup language." (Final Rejection at 5):

the usage log is transferred (see 219, FIG. 2).

Defining Audio Programming with HTML

Narrative text to be presented in the interactive, multimedia format made possible by the present invention may be advantageously expressed in the first instance using essentially conventional hypertext markup language, "HTML". FIG. 7 shows an example of the content of a portion of an illustrative HTML text file indicated generally at 450 used to create an audio file seen at 460 and a selections file indicated at 470.

The HTML file illustrated at 450 uses conventional <IMG> tags to identify image files, conventional emphasizing tag pairs <EM> and </EM> to designate highlighted passages, and conventional <A> and </A> HTML tag pairs to designate the anchor text and link target of a hypertext link. Utilizing conventional HTML to describe the narrative content to be presented in audio form provides several significant advantages, not the

least of which are:

conventional HTML composition software may be used to add the image and emphasis tags by means of visual tools which eliminate the need for hand-coding on a character level;

(a) a narrative text version of the audio programming may be viewed and printed, including both the emphasized text and the imbedded images, using most popular web browsers;

existing HTML files may be readily converted into audio multimedia presentations with little or no HTML editing being required;

HTML file may be made available from a server in a form which can be viewed in the normal way by any web browser yet and alternatively presented accordance with the invention in the form of an interactively browsable audio program with synchronized images;

the HTML file may be supplied along with the audio file as a transcript for the audio presentation, and to permit the audio presentation to be indexed and searched; and

the HTML may be automatically converted into the combination of an audio file using conventional speech synthesis techniques to process the narrative text with the HTML tags being used to compile a selections file which enables the player to interactively browse the audio file using highlighted and linked passages, and to synchronize the image presentation with the audio file.

(Logan et al., column 43, lines 15-60) Once more, the cited passage does not teach the substance Claim 1, including context based features, while adding the limitation "wherein the page being audibly read is a markup language page" and, for that reason, does not

anticipate Claim 11.

Applicants respectfully submit that Claim 11 of the claimed invention is not anticipated by the disclosure of Logan et al.

CLAIM GROUP 2

As noted above, the claims of Group 2 (claims 2 and 13) each recite the ability to access a different URL, and this feature is not required by the claims of Group 1. This feature underscores the multimodal browser context of the invention and the context based nature of the commands.

Claims 2 and 13

With regard to Claims 2 and 13, the Examiner found that "*Logan et al.* discloses the Program_Segments record URL field specifies the location file containing the program segment in the file storage facility 304 (column 17, line 62 to column 18, line 16; Figure 4); thus, the user listens to audio segments as stored resources based on URL[]s." (Final Rejection at 5) Applicants respectfully submit that the Examiner erred.

In finding Claims 2 and 13 to be anticipated by Logan et al., the Examiner has relied on Figure 4 from the disclosure of Logan et al. That figure, however, contemplates locating audio files over the Internet and playing them but does not anticipate "wherein the browser action comprises accessing a different Uniform Resource Locator." Nor does Figure 4 of Logan et al. require use of a browser as the means to access files over the Internet.

Just as Figure 4 of Logan et al. does not anticipate Claim 2 or Claim 13 of the claimed invention, the portion of the specification of Logan et al. cited by the Examiner do not anticipate the claims, either:

> The Program_Segment record's URL field specifies the location of the file containing the program segment in the file storage facility indicated at 304 in FIG. 4 (i.e., normally on the FTP server 125 seen in

> FIG. 1, but potentially including storage areas on the web server 141 or at any other accessible location on the Internet). In addition, the subscriber may wish to designate for future play a program segment already loaded into the player 103 by virtue of a prior download. The subscriber may elect to include an already loaded file because it was not reached in a prior playback session or because the subscriber wishes to replay the selection. In that event, the ProgramID of such a selection is nonetheless included in the uploaded selection list (Requested Table 301), recognizing that at the time of actual download, the player 103 will only request the transfer of those program segments not already present in local storage. The uploaded Requested list 301 should accordingly be understood to be indicative of the requested content of a future planned playback session and not necessarily a listing of programs to be downloaded. The selection of files to download is preferably made by the player which issues FTP download requests from the server by specifying the URLs of the needed files.

(Logan et al., column 17, line 62 – column 18, line 16) The cited passage does not anticipate Claim 2 or Claim 13 because it does not disclose the substance Claim 1 (or Claim 12) while adding that the browser action is comprised of accessing a different Uniform Resource Locator (URL) and rendering a page specified by the URL, as in Claims 2 and 13. Thus, the substance of dependent Claims 2 and 13 is not anticipated by the portion of the disclosure of Logan et al. cited by the Examiner in support of rejection.

ARGUMENT VIIID. REJECTIONS UNDER 35 U.S.C. §103

There are no rejections under 35 U.S.C. §103.

ARGUMENT VIIIE. REJECTION OTHER THAN 35 U.S.C. §§102, 103 AND 112

There are no rejections other than under 35 U.S.C. §§ 102, 103, and 112.

IX. APPENDIX OF CLAIMS INVOLVED IN THE APPEAL (37 C.F.R. §1.192(C)(9))

The text of the claims involved in this Appeal are:

1. A system for providing context based verbal commands to a multi-modal browser, comprising:

a context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user;

a store for storing a current context of the audio queue; and

a speech recognition engine for recognizing and registering voice commands, wherein said speech recognition engine compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison.

2. The system as recited in claim 1, wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and rendering a page specified by the URL.

3. The system as recited in claim 1, wherein when a first tag is used to designate the audio context, recognized voice commands associated with the audio context are ignored unless an audio context has been established, and wherein if a context has been established, a Uniform Resource Locator (URL) is followed after appending the current context.

4. The system as recited in claim 3, wherein said first tag is designated a REQUIRED tag.

5. The system as recited in claim 3, wherein when a second tag is used to designate the audio context, if a context is established, it is appended before driving the URL, and wherein if no context is established, the URL is followed without appending anything.

6. The system as recited in claim 5, wherein the second tag is designated an OPTIONAL tag.

7. The system as recited in claim 5, wherein when a third tag is used to designate the audio context, the context is not appended even if it is defined.

8. The system as recited in claim 7, wherein the third tag is designated an IGNORE tag.

9. The system as recited in claim 7, wherein when a fourth tag is used to designate the audio context, the command is driven only if a context is not defined.

10. The system as recited in claim 9, wherein the fourth tag is designated an INVALID tag.

11. The system as recited in claim 1, wherein the page being audibly read is a markup language page.

12. A computer implemented method for providing context based verbal commands to a multi-modal browser, comprising the steps of:

building a context based audio queue based on the contents of markup language page being audibly read by the multi-modal browser to a user;

storing a current context of the audio queue; and

recognizing and registering voice commands, wherein the current audio context is compared with a voice command, thereby causing the multi-modal browser to perform an action based on the comparison.

13. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 12, wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and displaying the contents of the

URL.

14. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 12, wherein when a first tag is used to designate the audio context, recognized voice commands associated with the audio context are ignored unless an audio context has been established, and wherein if a context has been established, a Uniform Resource Locator (URL) is followed after appending the current context.

15. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 14, wherein said first tag is designated a REQUIRED tag.

16. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 13, wherein when a second tag is used to designate the audio context, if a context is established, it is appended before following the URL, and wherein if no context is established, the URL is driven without appending anything.

17. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 16, wherein the second tag is designated an OPTIONAL tag.

18. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 16, wherein when a third tag is used to designate the audio context, the context is not appended even if it is defined.

19. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 18, wherein the third tag is designated an

IGNORE tag.

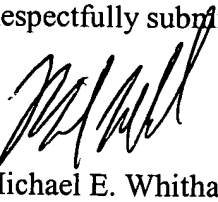
20. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 18, wherein when a fourth tag is used to designate the audio context, the command is driven only if a context is not defined.

21. The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 20, wherein the fourth tag is designated an INVALID tag.

X. Other Materials that Appellant Considers Necessary or Desirable

There are no other materials considered necessary or desirable for consideration in this appeal.

Respectfully submitted,



Michael E. Whitham
Registration No.32,635

Whitham, Curtis & Christofferson, P.C.
11491 Sunset Hills Road, Suite 340
Reston, VA 20190
Tel. (703) 787-9400
Fax. (703) 787-7557
Customer No. 30743